

INTERFACE FOR A UNIVERSAL SEARCH

BACKGROUND OF THE INVENTION

A. Field of the Invention

[0001] The present invention relates generally to database search engines, and more particularly, to user interfaces that are used to present the results of a search.

B. Description of the Related Art

[0002] The World Wide Web ("web") contains a vast amount of information. Search engines assist users in locating desired portions of this information by cataloging web documents. Typically, in response to a user's request, the search engine returns references to documents relevant to the request.

[0003] Search engines may base their determination of the user's interest on search terms (called a search query) entered by the user. The goal of the search engine is to identify links to high quality relevant results based on the search query. Typically, the search engine accomplishes this by matching the terms in the search query to a corpus of pre-stored web documents. Web documents that contain the user's search terms are considered "hits" and are returned to the user.

[0004] The corpus of pre-stored web documents may be stored by the search engine as an index of terms found in the web pages. Documents in the index may be categorized into one or more categories. Examples of such categories include general web documents, commercial products, images, sponsored

advertisements/links, and news documents. Dividing the index into categories can be beneficial in helping users find more relevant results to their search queries. For example, a user interested in finding a digital camera to purchase may be particularly interested in specific models and prices of digital cameras and not as interested in information relating to the taking of digital pictures.

[0005] One problem associated with search engines capable of performing multi-category searches is how to present the different category options to users. One known technique is to allow the user to pick the desired category and to then enter a search query. The user may, for example, first select a "news" tab from the search engine's home page and then enter a search query that will be applied to an index of news documents. Although this technique gives users a great deal of control in directing their search queries, a practical problem with this technique is that a large majority of users tend to ignore the category tabs, resulting in their search query being directed to the default category.

[0006] Thus, there is a need in the art to assist users in more effectively searching information segmented by category.

SUMMARY OF THE INVENTION

[0007] One aspect of the invention is directed to a method that includes performing a search to determine documents related to a search query. The search is performed on a database containing a number of document categories, and obtains a list of search results corresponding to each of at least two of the categories. The method further includes ranking the lists of search results

relative to one another and generating a document in which the search results are placed in an organization determined based, at least in part, on the ranking of the lists of search results.

[0008] Another aspect of the invention is directed to a search engine that includes a search component configured to determine documents related to a search query. The search is performed on a database containing a number of document categories and obtains a list of search results corresponding to each of at least two of the categories. A ranking component ranks the lists of search results relative to one another and an interface generation component generates a document in which the search results are placed in an organization determined based, at least in part, on the ranking of the lists of search results.

[0009] Yet another aspect of the invention is directed to a method of organizing documents categorized into a plurality of categories. The method comprising ranking the plurality of categories based on contents of the documents in each of the categories and generating a results document in which the documents are organized by category and in which the documents of higher ranking categories are more prominently placed in the results document.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate an embodiment of the invention and, together with the description, explain the invention. In the drawings,

[0011] Fig. 1 is an exemplary diagram of a network in which systems and methods consistent with the principles of the invention may be implemented;

[0012] Fig. 2 is an exemplary diagram of a client and/or server shown in Fig. 1;

[0013] Fig. 3 is a block diagram illustrating the classification of documents into different categories;

[0014] Fig. 4 is a block diagram conceptually illustrating functional components of the search engine shown in Fig. 1;

[0015] Fig. 5 is a diagram illustrating an exemplary web page that contains search results organized into a number of visual categories; and

[0016] Fig. 6 is a flow chart illustrating operations that may be performed consistent with aspects of the invention.

DETAILED DESCRIPTION

[0017] The following detailed description of the invention refers to the accompanying drawings. The detailed description does not limit the invention.

[0018] As described herein, a search engine returns search results to a user in which the search results are organized by document category. The different document categories are organized in the document returned to the user based, at least in part, on the relevance of the categories to the user search query.

EXEMPLARY NETWORK OVERVIEW

[0019] Fig. 1 is an exemplary diagram of a network 100 in which systems and methods consistent with the principles of the invention may be implemented.

Network 100 may include multiple clients 110 connected to one or more servers 120 via a network 140. Network 140 may include a local area network (LAN), a wide area network (WAN), a telephone network, such as the Public Switched Telephone Network (PSTN), an intranet, the Internet, or a combination of networks. Two clients 110 and one server 120 have been illustrated as connected to network 140 for simplicity. In practice, there may be more or fewer clients and servers. Also, in some instances, a client may perform the functions of a server and a server may perform the functions of a client.

[0020] Clients 110 may include client entities. An entity may be defined as a device, such as a wireless telephone, a personal computer, a personal digital assistant (PDA), a lap top, or another type of computation or communication device, a thread or process running on one of these devices, and/or an object executable by one of these device. Server 120 may include server entities that process, search, and/or maintain documents in a manner consistent with the principles of the invention. Clients 110 and server 120 may connect to network 140 via wired, wireless, and/or optical connections. Also, while server 120 is illustrated as a single entity, in practice, server 120 may be implemented as a number of networked server devices.

[0021] Clients 110 may include client software, such as browser software 115. Browser software 115 may include a web browser, such as the existing Microsoft

Internet Explorer or Netscape Navigator browsers. For example, when network 140 is the Internet, clients 110 may navigate the web via browsers 115 and, in particular, access server 120 via browsers 115.

[0022] Server 120 may operate as a web server and include search engine 125. In one implementation, search engine 125 may function as a query-based web page search engine. In response to client requests, search engine 125 may return sets of documents to clients 110. The documents may be returned to clients 110 as a web page containing a list of links to web pages that are relevant to the search query. This list of links may be ranked and displayed in an order based on the search engine's determination of relevance to the search query.

[0023] When performing a search for documents related to a user search query, search engine 125 may search database 127. Database 127 may include a number of documents, such as documents gathered from resources coupled to network 140 (e.g., web documents). The documents in database 127 may be segmented into categories, as will be described in more detail below.

[0024] A document, as the term is used herein, is to be broadly interpreted to include any machine-readable and machine-storable work product. A document may be an e-mail, a file, a combination of files, one or more files with embedded links to other files, a news group posting, a web advertisement, or the like. In the context of the Internet, a common document is a web page. Web pages often include textual information and may include embedded information (such as meta information, images, hyperlinks, etc.) and/or embedded instructions (such as Javascript, etc.).

EXEMPLARY CLIENT/SERVER ARCHITECTURE

[0025] Fig. 2 is an exemplary diagram of a client 110 and/or server 120 according to an implementation consistent with the principles of the invention. Client/server 110/120 may include a bus 210, a processor 220, a main memory 230, a read only memory (ROM) 240, a storage device 250, one or more input devices 260, one or more output devices 270, and a communication interface 280. Bus 210 may include one or more conductors that permit communication among the components of client/server 110/120.

[0026] Processor 220 may include one or more types of conventional processors or microprocessors that interpret and execute instructions. Main memory 230 may include a random access memory (RAM) or another type of dynamic storage device that stores information and instructions for execution by processor 220. ROM 240 may include a conventional ROM device or another type of static storage device that stores static information and instructions for use by processor 220. Storage device 250 may include a magnetic and/or optical recording medium and its corresponding drive.

[0027] Input device(s) 260 may include one or more conventional mechanisms that permit a user to input information to client/server 110/120, such as a keyboard, a mouse, a pen, voice recognition and/or biometric mechanisms, or the like. Output device(s) 270 may include one or more conventional mechanisms that output information to the user, including a display, a printer, a speaker, or the like. Communication interface 280 may include any transceiver-like mechanism that enables client 110/120 to communicate with other devices

and/or systems. For example, communication interface 280 may include mechanisms for communicating with another device or system via a network, such as network 140.

[0028] The software instructions defining search engine 125 and browser software 115 may be read into memory 230 from another computer-readable medium, such as data storage device 250, or from another device via communication interface 280. The software instructions contained in memory 230 cause processor 220 to perform processes that will be described later. Alternatively, hardwired circuitry may be used in place of or in combination with software instructions to implement processes consistent with the present invention. Thus, implementations consistent with the principles of the invention are not limited to any specific combination of hardware circuitry and software.

[0029] As mentioned, server software may implement a search engine 125 that, based on a user query, returns a web page containing a list of links to documents that the search engine considers to be relevant to the search query. The links presented to the user may be from multiple document categories and the different categories may be displayed in different sections of the web page (such as, for example, different visual sections). Consistent with an aspect of the invention, the choice of which links and categories to include may be based on relevance to the user search query.

SEARCH ENGINE 125

[0030] As previously mentioned, search engine 125 may operate on an index of documents stored in database 127. The documents in database 127 may include different categories of documents. In some implementations, each different category of documents may be stored as a separate document index. Fig. 3 is a block diagram illustrating classification of documents into different categories. Incoming documents, such as documents crawled from the web or otherwise input to search engine 125, may be input to a classifier 300. Classifier 300 may classify the documents based on their content. Possible classifications include sponsored links, news documents, product documents, documents summarizing discussion groups, images, and general web documents. Other classifications are, of course, possible. A product document may be a web page for a product that is for sale. A discussion group document may be an archive of an on-line discussion group or web pages that include the discussion group. A news document may include documents relating to news items or current events. Sponsored links may include links that a third-party pays the search engine operator to display (e.g., advertisements). In some implementations, a document may be classified as belonging to multiple classifications.

[0031] Classifier 300 may be an automated classifier that bases its decisions on, for example, the source of the documents, document content, and/or document structure. Techniques for classifying documents are well known in the art and will not be described further herein.

[0032] The classified documents may be stored as separate categories or “genres” within the complete document database 127, illustrated as sponsored links category 310, news category 311, products category 312, discussion groups category 313, images category 314, and general web documents category 315. One of ordinary skill in the art will recognize that more, fewer or different categories could be used. Each category 310-315 may be separately indexed by search engine 125.

[0033] Fig. 4 is a block diagram conceptually illustrating functional components of search engine 125. Search engine 125 may include a search component 401, a ranking component 402, and an interface generation component 403.

[0034] In general, search component 401 may receive user search queries, access database 127 based on the search queries, and obtain lists of links 410 (e.g., URLs) of relevant documents. The lists of links may also include information that generally attempts to describe the content of the web documents associated with the links. Each list in the lists of links may be associated with one of document categories 310-315. Thus, the first list may be of sponsored links that are relevant to the user search query, the second list may be of news articles relevant to the user search query, etc. The lists may be generated, for example, by search component 401 by individually applying the user search query to each document category 310-315.

[0035] Within each list of relevant links, the list may be ordered by perceived relevance to the user search query. In general, processes for searching an

indexed corpus of documents to return relevant documents are well known in the art. Accordingly, the functionality of search component 401 related to performing a search based on a user query and ranking the results within a list according to perceived relevance will not be described further herein.

[0036] The lists of links 410 may be ranked by ranking component 402.

Ranking component 402 may analyze the links in lists 410 and/or the documents corresponding to the links and, based on the analysis, determine a relative relevance of the lists. In other words, based on the search results generated from each category 310-315, ranking component 402 may rank the categories by perceived relevance to the user search query. Ranking component 402 may base its ranking on the documents corresponding to the links in each list and/or the user search query. For example, ranking component 402 may generally compare the search query to the contents of the documents in each list and base its ranking values on the closeness of the comparison. Consider the search query “buy athletic shoes.” For this search query, ranking component 402 may determine that the user is most likely interested in athletic shoes that are for sale. Accordingly, ranking component may rank the “products” category highly. The links in the list of links that correspond to the products category are likely to be links that correspond to web pages that are offering shoes for sale.

[0037] As another implementation of ranking component 402, the ranking component may look for terms in the search query that tend to indicate a particular category is likely to be related. For example, the word “buy” in a search query may indicate the “products” category should be ranked highly.

[0038] Interface generation component 403 may generate a document, such as a web page, that contains the results of the search query. The results may be organized on the web page in a manner that generally corresponds to the rankings generated by ranking component 402. Categories that are determined to be more relevant than others by ranking component 402 may be presented more prominently on the web page.

[0039] Fig. 5 is a diagram illustrating an exemplary web page 500 that contains search results organized into a number of categories. The results on web page 500 were returned by search component 401 for the exemplary search “foo fighters.” The various sections in web page 500 are demarcated by dashed lines. As illustrated, the largest section, section 510, contains general web page results for the search query “foo fighters.” Thus, general web page category 315 may have been selected as the highest ranked category by ranking component 402. In some implementations, because search engine users may expect to receive links derived from a general index of the web, interface generation component 403 may always construct web page 500 with the general web page category 315 as the most prominent category.

[0040] Section 510 may contain a number of links (e.g., hyper-text transfer protocol links) to relevant documents, such as link 520. Link 520 may additionally include a textual snippet 521 that provides additional information about the document referred to by link 520. Snippet 521 may be derived from the contents of the document.

[0041] Products section 511 and news section 512 are also relatively prominently displayed in web page 500. The links in products section 511 and news section 512 may therefore have been determined to be relevant to the user's search query, although potentially not as relevant as the links in general web results section 510. Accordingly, sections 511 and 512 may include fewer links than section 510. In particular, as shown, section 511 includes two links, links 525 and 526. Section 512 may also include two links. The links in products section 511 may be derived from products category 312 and may include information relevant to a user wishing to purchase a product, such as images of the products, prices, and the names of the merchants selling the products. In this case, the search term "foo fighters" is a musical group and the most relevant product links are two links selling one of the group's music albums. Products section 511 also includes a link 528 that indicates that additional product links are available. As shown, link 528 indicates how many other hits were generated in the products category. By selecting link 528, the user may be presented with another web page showing additional links from product category 312.

[0042] As shown, news section 512 may be structured similarly to products section 511. News section 512 contains a number of links to news articles from news documents category 311. News section 512 may also include a link indicating additional news items are available.

[0043] Other categories, such as image section 513 and groups section 514 may also be displayed in web page 500. These categories may correspond to categories that ranking component 402 determines to be less relevant than the

categories corresponding to sections 510, 511, and 512. As shown, no links may be explicitly shown in image section 513 and groups section 514. Instead, sections 513 and 514 may only indicate the number of hits in that category and a link to the results in that category. In this manner, the user is made aware that these categories are available and contain possibly relevant results, but web page 500 is not cluttered with an excessive amount of information.

[0044] It can be appreciated that the specific organization of sections 510-514 is exemplary. More particularly, the number, size, and layout of the sections could be changed.

[0045] Fig. 6 is a flow chart illustrating operations that may be performed by search engine 125 consistent with aspects of the invention. Search engine 125 may receive a search query from a user (act 601). Search component 401 may search multiple categories of documents, such as document categories 310-315, based on the search query (act 602). The search may result in, for at least two categories, a list of relevant documents or links to the relevant documents. Ranking component 402 may then rank the various categories based on the results returned from search component 401 (act 603). Interface component 403 may then generate the resultant web page that is to be returned to the user. As illustrated in Fig. 5, above, the web page may include links from multiple categories, organized, at least in part, based on category relevance (act 604). Server 120 may return the generated web page to the user (act 605).

CONCLUSION

[0046] By displaying a number of category sections in web page 500, search engine 125, as well as presenting a number of links from a primary category (such as the category that the search engine perceives to be most relevant to the user query), presents links from additional categories. Users may quickly scan links, and if the user realizes that they are more interested in a category other than the main category, the user can select a document from the other category or a link to additional results from the other category. In this manner, the user search experience can be improved without requiring the user to choose a category of interest before sending the search query to the search engine.

[0047] It will be apparent to one of ordinary skill in the art that aspects of the invention, as described above, may be implemented in many different forms of software, firmware, and hardware in the implementations illustrated in the figures. The actual software code or specialized control hardware used to implement aspects consistent with the present invention is not limiting of the present invention. Thus, the operation and behavior of the aspects were described without reference to the specific software code -- it being understood that a person of ordinary skill in the art would be able to design software and control hardware to implement the aspects based on the description herein.

[0048] The foregoing description of preferred embodiments of the present invention provides illustration and description, but is not intended to be exhaustive or to limit the invention to the precise form disclosed. Modifications and variations are possible in light of the above teachings or may be acquired

from practice of the invention. For example, although many of the operations described above were described in a particular order, many of the operations are amenable to being performed simultaneously or in different orders to still achieve the same or equivalent results.

[0049] No element, act, or instruction used in the description of the present application should be construed as critical or essential to the invention unless explicitly described as such. Also, as used herein, the article “a” is intended to include one or more items. Where only one item is intended, the term “one” or similar language is used.